Using apply() and charts

Introduction

Your task is to build a data frame that will contain information containing diabetes data in Northern Ireland. This data is fictitious, and any patterns found within it will be random.

Complete the following steps:

(1) Import the diabetes data available on Blackboard called diabetes-md into a new data frame called **diabetes_data**.

(2) Show the **structure** and **class** type of diabetes_data.

(3) We need to check whether there's any missing data. Display the missing values in the dataset. Decide what to do with missing values. What should you do with missing data? What effect will it have on remaining data? Choose what to do based on your analysis of missing values.

(4) The diabetes type and status attributes need to be refactored. The type attribute does not need to have an order. The status needs to have an order set as **poor** followed by **improved** and then **excellent**. Show the new structure of the data frame.

(5) Configure the names of the attributes to

- Patient name
- NI address
- Type
- Age
- Health status

(6) Create a copy of the names attribute into a data frame called **patient_names**. Show the first 10 names contained in **patient_names**.

(7) Display the class of each element in the data frame using the **lapply()** function.

(8) Remove the missing values from the **diabetes_data** data frame. Count and show the number of remaining records.

(9) Display 2 charts side-by-side. The first chart should show Health status with Age. The second chart should be a bar plot of Age. Label each chart with a title, and x,y axis labels.